NBER WORKING PAPER SERIES

FETAL SHOCK OR SELECTION? THE 1918 INFLUENZA PANDEMIC AND HUMAN CAPITAL DEVELOPMENT

Brian Beach
Joseph P. Ferrie
Martin H. Saavedra

Working Paper 24725
http://www.nber.org/papers/w24725

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018

We are grateful for feedback from Douglas Almond, Alan Barreca, Ryan Brown, Walker Hanlon, John Parman, Nick Sanders, Duncan Thomas, and Nic Ziebarth. We are also thank participants at the 2017 NBER Summer Institute Development of the American Economy meeting and the 2018 World Economic History Congress.The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Fetal Shock or Selection? The 1918 Influenza Pandemic and Human Capital Development
Brian Beach, Joseph P. Ferrie, and Martin H. Saavedra
NBER Working Paper No. 24725
June 2018, Revised September 2018
JEL No. I1,J0,N12

ABSTRACT

Almond (2006) argues that in utero exposure to the 1918 influenza pandemic lowered socioeconomic status in adulthood, whereas Brown & Thomas (2018) find that the effect disappears after controlling for parental characteristics of the 1919 birth cohort. We link microdata from the 1920 and 1930 censuses to WWII enlistment records and city-level in uenza data. This allows us to adopt an empirical approach that overcomes the selection concerns raised by Brown & Thomas (2018). Results indicate that in the absence of the pandemic, the 1919 birth cohort would have been more likely to graduate from high school, an effect that is largely unaffected by including parental controls and city-specific time trends. Adding household fixed effects (and thus exploiting variation among brothers) yields similar but somewhat larger results.

Brian Beach
Department of Economics
College of William and Mary
P.O. Box 8795
Williamsburg, VA 23187
and NBER
bbbeach@wm.edu

Joseph P. Ferrie
Department of Economics
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2600
and NBER
ferrie@northwestern.edu

Martin H. Saavedra
Department of Economics
Oberlin College
Oberlin, OH 44074
Martin.Saavedra@oberlin.edu

1 Introduction

In a highly influential paper, Almond (2006) argued that the 1918 influenza pandemic offers a natural experiment for testing Barker's 1995 hypothesis that in utero exposure to malnutrition increases susceptibility to heart disease in adulthood. The pandemic was an unexpected, severe, but ultimately temporary shock. As Almond noted, this allows the researcher to simply compare the adult outcomes of birth cohorts that were in utero during the pandemic to the outcomes of adjacent birth cohorts. Results from this specification indicate that in utero flu exposure not only decreased health in adulthood but it also lowered educational attainment, income, and socioeconomic status (SES) (Almond, 2006; Almond & Mazumder, 2005).

While Almond (2006) is often cited as evidence that early-life health insults can have lasting effects, recent work by Brown & Thomas (2018) offers an alternative explanation for these results. Specifically, the authors suggest that Almond's critical identifying assumption – that unobserved determinants of human capital vary continuously between birth cohorts – might not hold. As Brown & Thomas note, the 1918 pandemic coincided with the height of World War I (WWI) deployment and WWI veterans were positively selected from the overall population. Thus, the cohort that was in utero during the pandemic may have had worse outcomes in adulthood because they came from low SES families and not because of their exposure to flu while in utero. Importantly, Brown & Thomas (2018) show that Almond's core results are not robust to the inclusion of proxies for parental characteristics as controls.

We revisit Almond (2006) while taking the criticisms of Brown & Thomas (2018) seriously. Our efforts are made possible by Ancestry.com's recent digitization of the entire 1920 and 1930 federal censuses. These data allow us to construct an individual-level panel dataset by linking World War II enlistment records back to the census. This means that we can observe each individual twice: first as a child with their parents and again as an adult when they enlist.

Linking offers a number of advantages for this empirical setting. First, we do not have to rely on proxies for parental characteristics since we observe (and subsequently control for) own parental characteristics such as mother's and father's literacy, nativity, age, and occupational income score, as well as whether the family owned or rented their home, and the number of siblings. Second, and more importantly, linking allows us to adopt a more credible identification strategy. We assume that an individual's location at the time of enumeration is the same as their in utero environment. This allows us to exploit geographic variation in the intensity of the pandemic as an additional source of identification, which means that we can include birth year fixed effects in all of our specifications. These fixed effects account for the general rise in WWI deployment. Finally, the availability of these full count records also allows us to construct a large sample of brothers. For this subsample, we can include household fixed effects, which completely accounts for any parental selection.

We begin with a series of placebo tests to examine the credibility of our identification strategy. In contrast to Almond (2006), whose primary identifying assumption was that other determinants of human capital varied continuously across cohorts, our identifying assumption is that other determinants of human capital did not systematically vary with the intensity of the pandemic. Taking parental characteristics as outcome variables indicates that this assumption is likely valid, as we find no statistically significant relationships between any of our 10 measures of parental quality and intensity of exposure. Turning to long-run outcomes, we find robust evidence that in utero exposure to the pandemic reduced high school graduation rates. This is true before and after we control for parental characteristics, when we include city-specific time trends, and even when we analyze our household fixed effects model. As for biological outcomes, we find no consistent evidence that in utero exposure to the pandemic affected heights, weights, or BMI – the only proxies for health that are available in the enlistment records.

2 Data

2.1 Constructing a linked dataset

This section describes the construction of our linked dataset. We begin by identifying all male WWII enlistees whose reported year of birth falls between 1909 and 1922.[1] Parman (2015b) estimates that about 65 percent of men born between 1913 and 1923 enlisted in WWII, and so the sample is more representative of the male population than one might initially think.[2] The exception to this is that registrants could be rejected for failing to meet the minimum education or physical standards. This type of selection should bias against finding evidence that in utero flu exposure impaired human capital development.[3]

Our linking procedure builds upon our previous work (Long & Ferrie, 2013; Beach *et al.* , 2016) and can be summarized as follows: we take our set of enlistment records and cast a wide net to find a set of potential matches in the census immediately following their birth.[4] We then impose a number of requirements to ensure that the match is reasonable. If there is no match or if there is more than one potential match, we classify the enlistment record as unable to be linked.

[1]These records were digitized by the National Archives and Records Administration and are available at http://aad.archives.gov/aad/fielded-search.jsp?dt=893.

[2]Failing to link women is common in this literature, e.g. Aizer *et al.* (2016); Feigenbaum (2015); Long & Ferrie (2013). This is because women tend to change their name when they marry, and so without knowing their maiden name it is impossible to find their childhood records.

[3]A related source of selection is that the most exposed individuals may not survive until observation. This selection is present for both Almond (2006) and Brown & Thomas (2018) as well. This selection should also work against finding that in utero exposure impaired human capital development if we assume that these individuals would have had worse educational or health outcomes.

[4]For those reporting a year of birth between 1909 and 1919 this will be the 1920 census. For those reporting a year of birth between 1920 and 1922 the closest census will be the 1930 census. The 1920 census was enumerated in January and asks for an individual's age as of January 1, 1920. Thus, anyone born on or after January 1, 1920 should not appear in the 1920 census. It is worth pointing out that our ultimate empirical specification includes birth cohort fixed effects, which will address differential changes in the linking introduced by the fact that our 1920-1922 cohorts come from a different census. We prefer to include these cohorts for our baseline analysis because they received no direct exposure to the pandemic, and so they are our cleanest set of control cohorts. As a robustness check, we exclude these cohorts and we find nearly identical results.

We begin by standardizing all given names (e.g., "Ed" and "Eddie" would be recoded as "Edward") in both the enlistment records and the census records. A potential census match is one where the birth state is the same, the race is the same, the birth year is similar (plus or minus three years), the first initial of the standardized first names match, and the first initial of the last name matches. We then require that the raw first and last name strings are reasonably close. Specifically, we require that the Jaro-Winkler string distance for both the first names and the last names is greater than or equal to 0.8; a perfect match will have a string distance of 1. We do not require the names to match exactly for two reasons. First, during this time period census enumerators went door-to-door and recorded the information that was spoken to them; thus, the information recorded by the enumerator may represent a common spelling variant (e.g., "Elliot" and "Elliott"). Second, both the census records and the enlistment records were handwritten, and so relaxing the "exact name match" criteria allows us to deal with minor transcription errors.

A successful enlistment-to-census link is one where there is only one census record that survives this matching process.[5] Note however that the unique record may have an inconsistently reported birth year. The reason we allow year of birth to vary by up to three years is to accommodate the fact that the information comes from two different sources. From the census, we have to infer year of birth from reported age at the time of enumeration, which likely comes from a parent. The year of birth reported in the enlistment records, however, comes from the individual. If parents were less

[5]Recent work by Bailey *et al.* (2017) has raised concerns about false matches in linking studies. Our modified linking algorithm (relative to some of our previous work, e.g. Beach *et al.* (2016)) responds to these concerns in two ways. First, we cast a very wide net to increase the chance that the true link is in our set of potential matches. Second, if a record has more than one potential match, rather than try to break that tie by choosing the record with the closest name or closest age (as was done in our previous work) we simply classify the record as unable to be linked. To assess the false positive rate with our algorithm we take the full census sample and create a modified version where we take the original record and modify both the names and ages to incorporate the types of spelling errors, transcription errors, and misreporting of birth years outlined above and in Goeken *et al.* (2017). We then try to link from the original census to the modified census using our linking algorithm. Our linking algorithm yields a successful match rate of 36.5 percent and a false positive rate of 1.8 percent. This highlights the conservativeness of our algorithm.

numerate than their children then these numbers may not match. Another practical consideration is that the 1930 census was enumerated in April and collected age as of April 1, 1930. Thus, when we link to 1930, our inferred year of birth may be off by up to one year. Because accurately identifying birth year is essential for identifying in utero exposure, we further restrict our final linked sample to include only those whose year of birth matches is consistently reported. This procedure allows us to link 853,141 enlistees to a record in the 1920 or 1930 censuses.

2.2 Measuring in utero flu exposure

With our linked dataset in hand, we now turn our attention to measuring in utero flu exposure. Ideally we would have individual microdata identifying prenatal flu exposure (i.e., if and when exposure took place as well as a measure of severity). These data do not exist. Reliable data on maternal infection rates or morbidity are also unavailable as influenza morbidity data was not systematically collected prior to the pandemic. Beginning in 1900, however, annual influenza mortality data were collected and published in a systematic fashion for registration states and cities.[6] Thus, influenza mortality is a natural starting point for assessing in utero exposure.[7]

The most important feature of the influenza mortality data series, which was published alongside other important causes of death on an annual basis by the US

[6]Registration states and cities are those with laws requiring that mortality statistics be collected. In contrast to England, which standardized and mandated the reporting of deaths in 1846, the United States left this decision to state and local governments. Several large cities and states passed mandatory reporting laws by 1900, and in that year the Census Bureau worked with those registration areas to establish uniform reporting standards. The result of this was the adoption of a standardized death certificate and the international classification standard, as well as the distribution of "The Manual of International Classification of Causes of Death", which cross referenced terms appearing in causes of death from 1890 and 1900 reports with the new uniform classification standard.

[7]Whether influenza mortality tracks influenza morbidity trends reasonably well is a natural question. During the fall of 1918, the Public Health Service went door-to-door collecting morbidity and mortality information for 12 cities. Appendix Figure A.1, which appears in their report, plots weekly morbidity and mortality rates for five cities during the fall of 1918. The figure indicates the fatality rate tracks the case rate reasonably well on a weekly basis but with a bit of a lag. Given that we ultimately draw on annual-level data, this lag is unlikely to be much of a concern.

Census Bureau, is that the underlying data conform to a common reporting standard. This ensures that we are able to measure flu fatalities with a great degree of precision, but it does come with one drawback: the data are not comprehensive. Cities and states were only included in the published reports if the underlying data were deemed reliable. In 1900 there were 330 registration cities systematically collecting mortality data, but by 1920 the registration area included 662 cities spanning 41 states.

While influenza mortality data capture total influenza deaths in 1918, one concern is that influenza mortality will capture more than just the severity of the pandemic. For instance, early public health scholars often noted that clean water interventions lowered mortality from waterborne causes as well as causes that are not typically thought of as waterborne (e.g., influenza, tuberculosis, pneumonia, kidney failure, and heart failure).[8] Relatedly, Clay *et al.* (2015), document that mortality rates in 1918 were higher in places with more coal pollution and places with worse water quality. These relationships are attributable to the fact that air pollution and typhoid fever compromise an individual's immune system, making them more susceptible to influenza. In light of this, observing high influenza mortality rates in 1918 could mean that a city was hit relatively hard by the pandemic, or that a city had relatively worse water and air quality, or both. Since air pollution and water quality have also been shown to impair human capital development, influenza mortality rates may capture more than in utero flu exposure.[9]

Our solution to this problem is as follows. First, we generate a counterfactual estimate of influenza mortality in 1918. To do so, we transcribe all city-level mortality statistics spanning 1900-1930 from the annual Mortality Statistics reports. We then run a series of city-level regressions where we restrict the sample to the 1900-1917 period and regress ln(influenza deaths) on a city-specific linear time trend.[10] Taking

[8]This phenomenon is often referred to as the Mills-Reincke Phenomenon. See Ferrie & Troesken (2008) for more discussion as well as an empirical test of this theory in Chicago.

[9]See Sanders (2012) and Isen *et al.* (2017) on early-life exposure to air pollution and Beach *et al.* (2016) on early-life exposure to typhoid fever.

[10]We only run these regressions for the 287 cities that appear in every report.

the exponential of the predicted values from this regression yields a prediction of influenza fatalities in the absence of the pandemic for post-1917 years.[11] This allows us to construct a measure of excess influenza deaths by simply subtracting predicted influenza deaths in 1918 from actual influenza deaths in 1918.

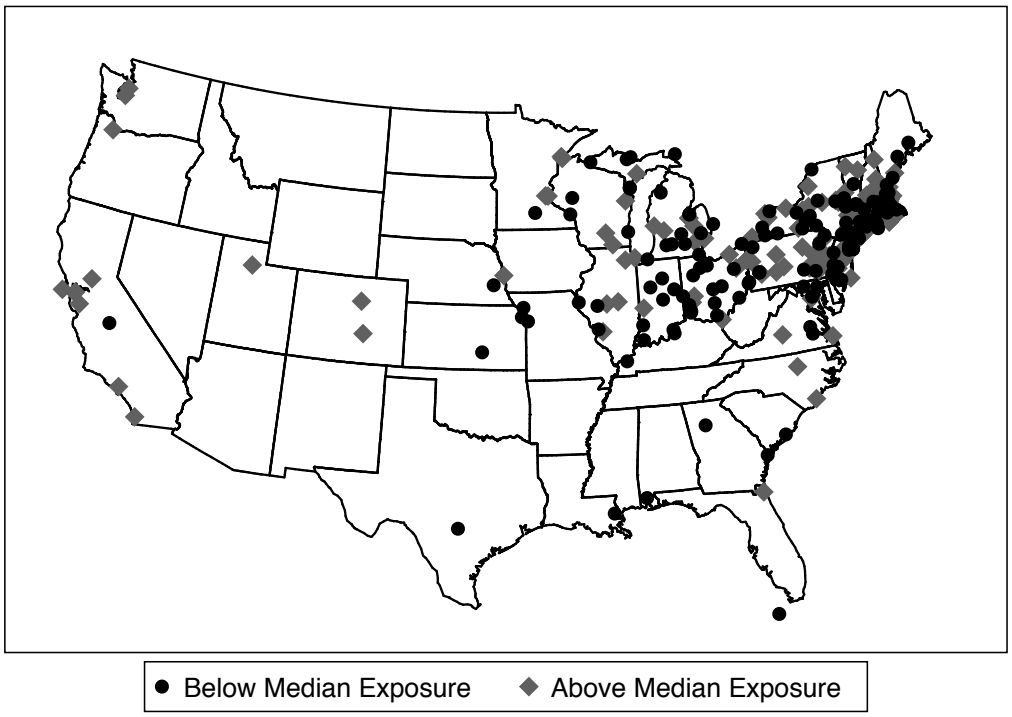
This measure of excess deaths gives us the unanticipated increase in influenza mortality due to the pandemic. Of course, we need to normalize this measure because excess deaths will be related to city size, and the goal is to have a measure of in utero exposure to influenza, not in utero exposure to large cities. Our options are to divide by population or to divide by predicted influenza deaths. Population and predicted influenza deaths are positively correlated: all else equal, a larger city should have more predicted influenza deaths. However, dividing by population ignores the fact that cities of similar sizes may have different underlying disease and pollution environments. Because of this, normalizing by population will likely overstate flu intensity for cities with particularly bad health environments. For this reason, and also because accurate population data are only available in census years, we use predicted influenza deaths as our denominator. Mechanically this measure is simply the ratio of the number of excess influenza deaths occurring in 1918 to the number of expected influenza deaths in 1918, where that expectation captures underlying trends in population growth and intrinsic differences in disease and pollution environments.

Appendix Figure A.2 plots the average excess influenza ratio by year. The figure spans 1900-1917 (the sample period) as well as 1918 through 1930 (our out of sample predictions).[12] There we see the severe and temporary nature of the 1918 pandemic as well as the overall fit of our model. Excess influenza remains close to zero until 1918, during which approximately 35 influenza deaths occurred for every expected influenza death. Influenza deaths are higher during than the 1920s than would have

[11] The natural logarithm ensures that predicted influenza deaths is always greater than zero.

[12] Alternatively, we could have used data from the 1900-1917 and 1920-1930 years, omitting the years during which the pandemic occurred. However, if the pandemic had lingering effects on influenza rates or city population counts, then data from the 1920-1930 period may be endogenous.

Figure 1: Spatial Variation in Flu Intensity



Notes: High exposure cities had an excess flu ratio greater than 28.2.

been predicted using the pre-pandemic data, however, even 13 years after 1917, excess influenza deaths are not far above zero. It appears influenza reached its new steady state in 1921.

Figure 1 provides an illustration of the spatial variation in our sample. We plot separate markers for cities that had above median vs below median exposure to the pandemic. The median excess flu ratio is 28.2. The key takeaway from Figure 1 is that there is meaningful sub-state variation in our measure of flu intensity: there are many situations where neighboring cities had different exposure to the pandemic.[13] This level of spatial variation, which has not been exploited in prior studies, will be particularly useful for examining the long-run effects of the pandemic.

[13]Appendix Figure A.3 plots a continuous version of this figure.

Combining our linked dataset with our measure of excess influenza yields a final sample of 218,662 linked records. The reduction in sample size largely stems from the fact that many of our linked individuals did not live in a city for which we can calculate exposure to the pandemic.[14] Summary statistics are reported in Appendix Table A.1. The average years of schooling for our sample is 11.4 with about 54 percent of our sample completing high school. In terms of family characteristics, the literacy rates among mothers and fathers were quite high at 94 and 92 percent, respectively. About 36 percent of families owned their homes and only 4.5 percent of mothers were in the labor force at the time of enumeration.

3 Empirical Approach

We adopt a difference-in-differences strategy to identify the impact of in utero flu exposure on adult outcomes. Specifically we estimate variations of two regressions. For the first, we regress:

$$y_{ibc} = \alpha_0 + \beta_b + \gamma_c + \delta \mathbf{1}\left[\text{yob} = 1919\right] \times \text{High}_c + \Gamma X_i' + \epsilon_{ibc} \tag{1}$$

where y_{ibc} is outcome y of individual i from birth year b in birth city c. The parameters β_b and γ_c are birth year and birth city fixed effects, respectively. The variable High_c is equal to one if city c had above median excess influenza in 1918. Our baseline specification restricts our analysis to a narrow window (cohorts born in 1918, 1919, or 1920) to better isolate the effect of in utero exposure. The vector X_i includes additional controls such as race and parental characteristics. For some specifications,

[14] We are unable to use 597,486 of our original links because we do not have flu data for their birth city. We throw out an additional 26,556 links because the child is not observed with both parents at the time of enumeration and we throw out an additional 4,296 links because one or more of the relevant parental controls is missing. Next we discard an additional 6,141 links because the parental age at time of birth is implausible (e.g., mother or father's age at the time of birth is over the age of 50 or under the age of 18). We discard these observations because we are concerned that the indicated mother or father is not actually the birth mother or birth father.

we only estimate a treatment effect for the 1919 birth cohort and include birth-city linear time trends. Standard errors are clustered at the birth city level.

Figure 2 depicts this identification strategy (and one of our main results). Specifically, we plot the average high school graduate rate by birth cohort for cities that will ultimately have above and below median exposure to the pandemic. To ease interpretation, we normalize high school graduation rates at the city-level. The figure reveals that for the 1909 to 1916 birth cohorts, graduation rates in cities with above and below median exposure tended to follow a nearly identical trend. Beginning with the 1917 cohort the average graduation rates start to fall relative to previous trends for both groups and the effects appear to be much more severe for those born in cities with above average exposure to the pandemic. While the 1919 cohort experiences the largest deviation from trend, by 1920 graduation rates appear to return to normal.

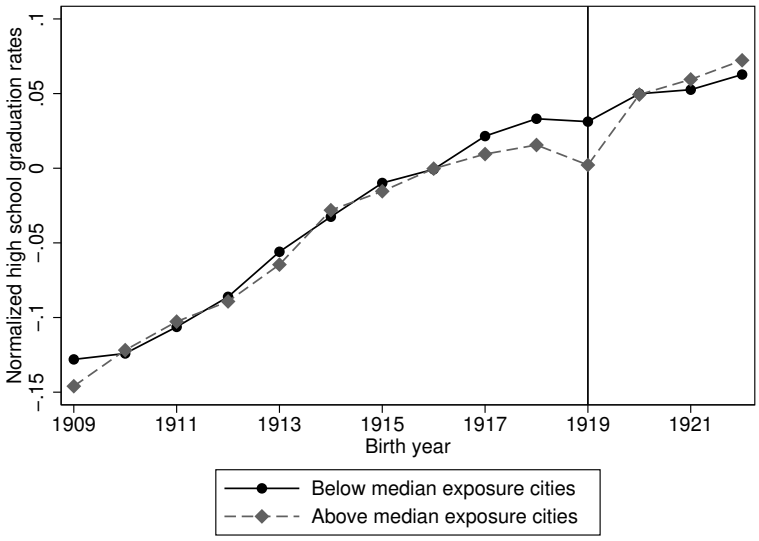
Figure 2 indicates that this difference-in-difference strategy is likely to satisfy the parallel trends assumption. However, it is worth noting that the slight deviations in trend for the 1917-1919 cohorts in low exposure cities may be due to early life pandemic exposure. A large literature has shown that early life exposure to disease and deprivation also impairs human capital development (Currie & Almond (2011)). Our regressions, however, treat those deviations as the counterfactual trend that would have been observed in high exposure cities if the pandemic had never occurred. To the extent that the stagnation in low exposure (i.e. control) cities is due to the pandemic, our estimates can be thought of as a lower bound.

While the previous specification assigns above median excess influenza cities as the treatment group and below median excess influenza cities as the control group, our second specification uses continuous variation in treatment status:

$$y_{ibc} = \alpha_0 + \beta_b + \gamma_c + \delta \mathbf{1}\left[\text{yob} = 1919\right] \times \frac{\text{flu}_{1918,c}}{E(\text{flu}_{1918,c}|\text{yob} = 1919)} + \Gamma X_i' + \epsilon_{ibc}. \quad (2)$$

The variable $\text{flu}_{1918,c}$ is excess influenza in 1918 for city c. This variable is divided by

Figure 2: Normalized graduation rates by birth cohort in above and below median pandemic exposure cities



Notes: Sample consists of 218,662 enlistees linked to either the 1920 or 1930 censuses. Graduation rates normalized by removing city-specific mean before collapsing to the cohort level. The median pandemic city had an excess flu ratio of 28.2.

$E(\text{flu}_{1918,c}|\text{yob} = 1919)$ so that δ is the average effect of the pandemic for the 1919 birth cohort. All other parameters are defined as above.

4 Results

Before proceeding to our main results, Table 1 provides additional support for our identification strategy. Each column corresponds to a different parental characteristic: father's age when the child was born, father's occupational income score, whether the father is literate, whether the father is foreign born, mother's age when the child was born, whether the mother is employed, whether the mother is literate, whether the mother is foreign born, whether the family owns their home, and number of siblings. All of these characteristics are available from the census where we observe the child with his parents. We take each of these characteristics as outcome variables and

estimate equations 1 and 2. The coefficients of interest from those specifications are the interaction between being born in 1919 and pandemic intensity.

None of the 20 regressions presented in Table 1 are statistically significant and many of the coefficients are close to zero. This is reassuring as it suggests that the parental selection concerns raised by Brown & Thomas (2018) are not also systematically related to the intensity of the pandemic. In other words, an empirical approach that includes cohort fixed effects and instead relies on variation in pandemic intensity as a source of variation is likely to lend itself to a causal interpretation.

Having established the credibility of our identification strategy, we now turn our focus to long-run outcomes. Our main results are presented in Figure 3. The top panels consider high school graduation as the outcome of interest while the bottom panels consider height at the time of enlistment. These are the main economic and biological outcomes that are available in the enlistment records. The left panels use the discrete treatment from equation 1 while the right panels use the continuous exposure from equation 2.

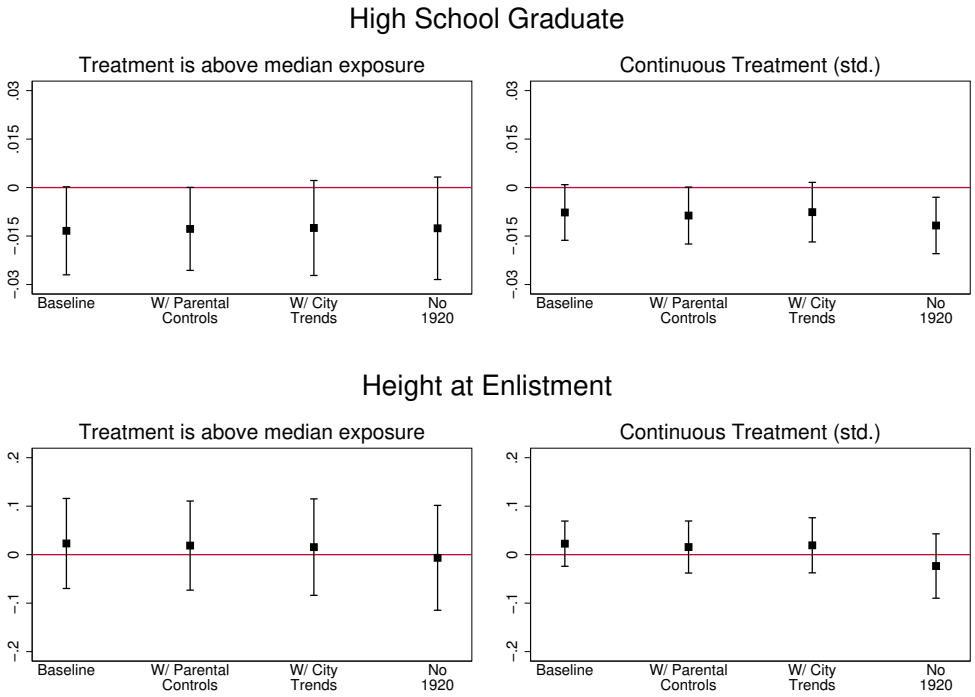
Within each panel we present results from four specifications. The baseline specifications correspond directly to equations 1 and 2. The "W/Parental Controls" specifications add all of the family characteristics that we examined in Table 1. The "W/City Time Trends" specifications include both the family characteristic controls and city-specific linear time trends. Finally, the "No 1920" specifications return to the baseline but omit the 1920 cohort, such that the analysis simply compares the 1919 birth cohort against the 1918 birth cohort. We omit the 1920 cohort for two reasons. The first is to ease any concerns about the 1920 cohort being observed in the 1930 census while the 1918 and 1919 birth cohorts are observed in the 1920 census. The second is because omitting the 1920 cohort offers the most conservative test of the Fetal Origins Hypothesis, as the implicit counterfactual to in utero exposure is exposure during infancy. Thus, this specification allows us to ask whether in utero exposure matters more than exposure in the post-natal period.

Table 1: Placebo Tests – Intensity of in utero exposure and family characteristics

Dependent Variable:	Father's age	Father's OccScore	Literate Father	Foreign Father	Mother's age
	(1)	(2)	(3)	(4)	(5)
Treatment is above median exposure	0.058	0.147	-0.003	-0.003	0.118
	(0.112)	(0.272)	(0.004)	(0.008)	(0.095)
Continuous treatment (std.)	0.068	0.131	0.001	0.001	0.141
	(0.074)	(0.174)	(0.003)	(0.005)	(0.086)
Dependent Variable:	Mother Works	Literate Mother	Foreign Mother	Fam. Owns Home	Num. Siblings
	(6)	(7)	(8)	(9)	(10)
Treatment is above median exposure	-0.004	0.000	-0.006	-0.001	0.025
	(0.003)	(0.004)	(0.007)	(0.008)	(0.035)
Continuous treatment (std.)	-0.003	-0.002	0.003	-0.001	0.008
	(0.003)	(0.003)	(0.006)	(0.006)	(0.026)

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (clustered at the city level) in parentheses. Each entry is the coefficient obtained from regressing the indicated outcome variable on the defined measure of flu intensity as well as cohort fixed effects and birth city fixed effects. Each regression includes 71,026 observations, corresponding to the linked birth cohorts born in 1918, 1919, or 1920. The above median exposure specifications interact an indicator for being born in 1919 with another indicator for being born in a city that had above median pandemic exposure (as measured by the excess influenza rate). The continuous treatment specifications interact each city's excess influenza rate with an indicator for being born in 1919. We standardize the excess influenza rate by dividing by the mean excess influenza rate, such that the coefficient can be interpreted as the pandemic's average impact.

Figure 3: Flu intensity and long-run outcomes



Notes: Sample consists of 71,026 individuals when the outcomes are high school completion and 42,424 observations when the outcome is height. Each regression includes cohort fixed effects and city fixed effects. The "W/Parental Controls" specifications add all of the family characteristics that we examined in Table 1. The "W/City Time Trends" specifications include both the family characteristic controls but also add in city-specific linear time trends. High exposure cities had an excess flu ratio greater than 28.2.

The results in Figure 3 are remarkably stable. We consistently find that in utero exposure decreased the likelihood of completing high school by about 1 to 1.5 percentage points. Each of these estimates is statistically significant at roughly the 5% level. For height, however, we find statistically insignificant effects. Moreover, the 95% confidence interval is not economically meaningful, as it suggests that in utero exposure may have affected height at the time of enlistment by up to one-tenth of an inch. Given the lack of statistical and practical significance, we interpret this as evidence that the pandemic did not meaningfully affect the height of enlistees.

5 Evidence from a wider set of cohorts

The previous section provided results from a narrow band of birth cohorts. As mentioned previously, these results are somewhat conservative because one of the two control cohorts (those born in 1918) are treated in the sense that they were exposed to the pandemic during infancy. In contrast to the results in the previous section, which only consider the 1918, 1919, and 1920 birth cohorts, the results in this section will consider every cohort from 1909 through 1922. The advantage of this expansion is that it allows us to construct a sufficiently large sample of brothers such that we can eventually include household fixed effects.

These results are presented in Table 2. Panel A presents results for high school graduation, Panel B presents results for total years of schooling, and Panel C presents results for height. In each panel, the first row models exposure as a simple indicator for being born in a high exposure city in 1919 and the second row models in utero exposure by interacting each city's excess influenza rate with an indicator for being born in 1919. Column 1 presents our baseline results. In column 2 we add controls for our parental characteristics (described above), while in column 3 we add city-specific linear trends. In column 4 we focus on brothers and add household fixed effects.[15]

[15]Since the 1909 to 1919 cohorts are observed in the 1920 census and the 1920 to 1922 cohorts

We find robust evidence that the pandemic decreased high school graduation rates by about 2 percentage points. This result survives the inclusion of household fixed effects when we model in utero exposure continuously, but the p-value falls just outside the 10-percent level of significance when those with below median exposure to the pandemic are considered as part of the control group. For schooling, we find consistent and negative effects, but we lose statistical significance in the household fixed effects specification. In the Appendix we show that omitting the 1930 links (the 1920 to 1922 birth cohorts) has little effect on our estimates.

Turning to biological outcomes, we find no effect on heights when influenza is measured discretely and, as before, the estimated coefficient is close to zero. When influenza exposure is measured continuously, the pandemic appears to have increased heights but this result is not robust to the inclusion of household fixed effects. Even for the specifications that are statistically significant, however, the estimates are not economically significant. The estimates suggests that the pandemic may have increased heights by less than one twentieth of one inch.[16] In the appendix we consider weight and BMI at the time of enlistment and find no evidence that these outcomes were affected by in utero exposure to the pandemic. When we consider the totality of these results, we conclude that there is little consistent evidence that the pandemic impaired health as measured at the time of enlistment. Of course, it is worth pointing out that these results should be interpreted with the caveat that we are observing enlistees in their 20s and 30s. Thus, it is possible that these results are driven by the fact that unfit individuals will not appear in our sample or that many of the health effects (e.g., heart disease) will not manifest for another 10 to 20 years or more.

are observed in the 1930 census, we won't observe any of the 1920 to 1922 birth cohorts with their older brothers born before 1920. To remedy this we go back to census manuscripts and identify any older brothers that are still in the household in 1930. If any of those older brothers happen to be successfully linked between the enlistment records and 1920, then we are able to generate a household identifier that is consistent regardless of whether the brother was born in 1909-1919 (and thus enumerated in 1920) or born between 1920 and 1922 (and thus enumerated in 1930).

[16]To put this in perspective, Parman (2015a) relates enlistee height to state and city-level disease environments and finds that a standard deviation increase in infant mortality rates decreases height by 0.3 inches, nearly 6 times greater than the effect we find.

Table 2: In utero exposure and long-run outcomes

	(1)	(2)	(3)	(4)
Panel A: DV is High School Graduate				
Treatment is above median exposure	-0.023***	-0.020***	-0.020***	-0.041
	(0.007)	(0.007)	(0.006)	(0.030)
Continuous treatment (std.)	-0.016***	-0.014***	-0.013***	-0.043*
	(0.004)	(0.005)	(0.004)	(0.022)
Observations	218,662	218,662	218,662	23,395
Panel B: DV is Years of Schooling				
Treatment is above median exposure	-0.087**	-0.070**	-0.050*	-0.051
	(0.034)	(0.030)	(0.028)	(0.141)
Continuous treatment (std.)	-0.063***	-0.055***	-0.037**	-0.069
	(0.021)	(0.018)	(0.018)	(0.095)
Observations	218,662	218,662	218,662	23,395
Panel C: DV is Height at Enlistment				
Treatment is above median exposure	0.013	0.015	0.027	-0.097
	(0.041)	(0.039)	(0.040)	(0.220)
Continuous treatment (std.)	0.044**	0.045**	0.047**	0.022
	(0.022)	(0.019)	(0.020)	(0.123)
Observations	168,317	168,317	168,317	18,051
Parental controls	N	Y	Y	Y
City-year trends	N	N	Y	Y
Household fixed effects	N	N	N	Y

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (clustered at the city level) in parentheses. Parental controls include: father's age at time of birth, mother's age at time of birth, father's OccScore, indicators for whether the father and mother can read and write, indicators for whether the father and mother are foreign born, an indicator for whether the mother was in the labor force at the time of enumeration, an indicator for whether the family owns their home, and number of siblings. The continuous treatment specifications interact each city's excess influenza rate with an indicator for being born in 1919. We standardize the excess influenza rate by dividing by the mean excess influenza rate, such that the coefficient can be interpreted as the pandemic's average impact.

6 Discussion

Overall the results in the previous two sections are remarkably consistent. The ability to include household fixed effects increases our confidence that the previous results are not due to parental selection but are instead attributable to exposure to the pandemic. In short, we conclude that in utero exposure to the 1918 influenza pandemic did impair human capital development. One natural question, however, is why we find evidence that is consistent with Almond (2006) while Brown & Thomas (2018) find that Almond's results are not robust to the inclusion of proxies of parental characteristics as controls.

There are many differences between our sample and the samples analyzed by Almond (2006) and Brown & Thomas (2018). We examine a linked sample of males that resided with both parents at the time of enumeration and ultimately enlisted in WWII. Moreover, due to our desire to exploit variation in pandemic intensity, we further require that our individuals resided in a city at the time of enumeration. Because we examine outcomes at the time of enlistment, our outcome variables are measured anywhere from 20 to 40 years earlier than in Almond (2006) and Brown & Thomas (2018).

In light of these differences, it is perhaps useful to try and draw a closer comparison between our results and the results that appear in Almond (2006) and Brown & Thomas (2018). Section VI of Almond (2006) asks whether there is evidence of a dose response. To do so, he constructs a proxy for state-level maternal infection rates and runs a series of regressions that are otherwise quite similar to our main results (i.e. restricting to the 1918, 1919, and 1920 birth cohorts).[17] The primary

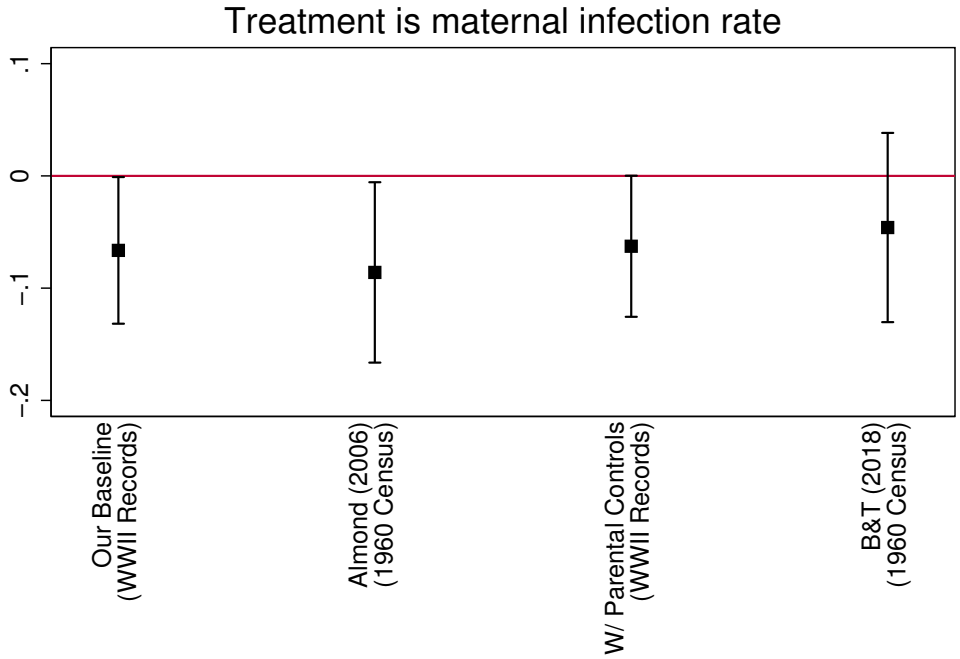
[17]Almond defines the maternal infection rate for the 1918, 1919, and 1920 cohorts as follows: $MIR_{s,1918} = 0$; $MIR_{s,1919} = \frac{MMR_{s,1918} - MMR_{s,1917}}{\kappa - MMR_{s,1917}}$; and $MIR_{s,1920} = \frac{MMR_{s,1919} - MMR_{s,1917}}{\kappa - MMR_{s,1917}}$, where $MIR_{s,t}$ denotes the maternal infection rate for cohorts born in state s during year t, $MMR_{s,t}$ denotes that maternal mortality rate in state s during year t (the year in which the child was in utero); and κ denotes maternal mortality conditional on influenza infection (1.4%).

limitation of this proxy relative to our preferred proxy is that it only exploits state-level variation, and thus does not fully exploit geographic variation in pandemic exposure. A second issue is that this proxy does not incorporate actual influenza mortality data. Nevertheless, when Brown & Thomas (2018) revisit this analysis they show that including their proxies for parental characteristics as controls attenuates Almond's findings, but notably unlike earlier specifications the inclusion of these controls does not flip the sign of the coefficient.

We reconstruct this maternal infection rate proxy, which allows us to draw comparisons between our findings and the findings in Almond (2006) and Brown & Thomas (2018). We present these results graphically in Figure 4. The baseline specification corresponds to equation 2 in that it includes cohort fixed effects and city fixed effects, but instead of continuous excess influenza as our treatment we use the maternal infection rate proxy described above. The next specification – "Almond" – corresponds to Brown & Thomas's replication of Almond, using the 5% sample from the 1960 census. Column 3 – "W/Parental Controls" – returns to our baseline specification but includes each of our 10 family background characteristics. Column 4 – "Brown and Thomas" – corresponds to the estimates that Brown & Thomas produce when they include their proxies for parental characteristics. As is clear from this figure, results are consistent across all four specifications: it is not possible to reject equality for any of these coefficients at standard levels of significance.

We interpret Figure 4 as suggestive evidence that our findings are not driven by the differences in samples. Appendix Table A.4 presents additional evidence in favor of this conclusion. There we apply the maternal infection rate approach to the 1918-1920 birth cohorts as observed in the 100% files of the 1940 census. The 1940 offers a unique set of advantages and disadvantages. The advantage of the 1940 census is that we are able to observe all enumerated individuals, which alleviates concerns about the bias of our linked sample. Further, roughly half of the individuals born between 1918 and 1920 (and born in states for which we can compute the maternal infection

Figure 4: Flu intensity and high school graduation



Notes: The baseline specification and "W/Parental Controls" specification analyze 1918, 1919, and 1920 birth cohorts from our linked sample. Sample is further restricted to those residing in a state for which maternal mortality data are available. Each of these regressions includes cohort fixed effects and city fixed effects. The "W/Parental Controls" specifications add all of the family characteristics that we examined in Table 1.This sample includes 53,214 linked individuals. Maternal infection rate is calculated as in Almond (2006). The "Almond" and "Brown and Thomas" specifications are recreated from Table 7 of Brown & Thomas (2018).

rate) are observed with both of their parents. This means that the 1940 census offers a rare opportunity to observe adult outcomes and parental characteristics at the same time. The fundamental drawback of the 1940 census, however, is that the reference date for the age question is April 1, 1940, which means that year of birth is measured much less precisely.

Appendix Table A.4 presents two sets of results. Panel A examines the impact of the in utero exposure on high school graduation rates. Column 1 examines the full census, while column 2 restricts the sample to the set of individuals that still reside at home. For those individuals, column 3 includes a host of parental controls. In all three specifications we see that in utero exposure decreased the likelihood that the individual would graduate high school. The point estimates are roughly half of the size of our baseline effect in the enlistment sample, perhaps reflecting that we are measuring in utero exposure with less precision.

Panel B of Appendix Table A.4 applies the same empirical specification but considers 8 different parental characteristics as outcome variables (whether the individual's mother or father graduated high school, mother and father's age when the child was born, whether the individual's mom was foreign born, whether the individual's parents were employed, and the father's occupational income score). Most of the coefficients are not economically meaningful, and only "mother is foreign born" is statistically significant at conventional levels. This panel suggests that this empirical approach is in fact a credible identification strategy.

In light of these results, it is an open question as to why Brown & Thomas (2018) find that including proxies for parental controls reverses Almond's original findings. One potential explanation may stem from the fact that Brown & Thomas (2018) are forced to use proxies for parental characteristics, which can be problematic if those proxies are systematically mismeasured.

To explore this further, let us consider a simple human capital model where son i's educational attainment (y_i^{son}) depends on in utero influenza exposure (flu_i), his

father's educational attainment (y_i^{father}), and exogenous unobserved characteristics (ϵ_i), which are independent of the other explanatory variables, so that:

$$y_i^{\text{son}} = \beta \text{flu}_i + \gamma y_i^{\text{father}} + \epsilon_i. \tag{3}$$

Further assume that as in Brown & Thomas and Almond, flu_i is an indicator variable that affects all children from the 1919 birth cohort and otherwise equals zero. Estimation of equation 3 would yield consistent estimates of β if y_i^{son}, flu_i, and y_i^{father} were observable.

The advantage of our linked data approach is that we are able to observe all three of these variables. Brown & Thomas, however, do not observe y_i^{father}. Instead, Brown & Thomas observe average parental characteristics for a state-year cell. Assuming the observed averages correspond to the truth, we can express these parental characteristics as $\bar{y}_{s,t}^{\text{father}} = y_i^{\text{father}} + \nu_i$, where ν_i is the deviation from the mean for father i and is by assumption independent of $\bar{y}_{s,t}^{\text{father}}$.[18] Since flu_i is only a function of t, it follows that ν_i is also independent of flu_i. Then estimation of

$$y_i^{\text{son}} = \beta \text{flu}_i + \gamma \bar{y}_{s,t}^{\text{father}} + u_i \tag{4}$$

yields consistent estimates since $u_i = -\gamma\nu_i + \epsilon_i$, both of which are exogenous to the explanatory variables. Thus, if parental characteristics are accurately measured, Brown & Thomas's inclusion of proxies would successfully account for parental characteristics and return causal estimates of the effects of the 1918 influenza pandemic.

Now suppose that $\bar{y}_{s,t}^{\text{father}}$ is mismeasured. This mismeasurement could be due to age heaping: the phenomenon that relatively uneducated parents tend to round the ages of their children (e.g. age 10 instead of 9 or 11), whereas relatively educated parents correctly report the ages of their children. This phenomenon would cause

[18] The two are by definition uncorrelated, but we strengthen the assumption to independence to simplify the discussion.

average parental characteristics to be biased downwards for heaped ages and upwards for non-age heaped ages. Note that Brown & Thomas obtain parental characteristics from the 1930 census, where the cohort that was in utero during the pandemic should be individuals with a reported age of 10.

Formally, let $\tilde{\bar{y}}_{s,t}^{\text{father}}$ be the mismeasured average parental characteristics, and $e_{s,t}$ be the error, so that $\tilde{\bar{y}}_{s,t}^{father} = \bar{y}_{s,t}^{\text{father}} + e_{s,t}$. For heaped ages, $e_{s,t} < 0$; for non-heaped ages, $e_{s,t} > 0$.

Consider the estimation of

$$y_{i,s,t}^{\text{son}} = \beta \text{flu}_i + \gamma \tilde{\bar{y}}_{s,t}^{\text{father}} + \eta_i. \tag{5}$$

The error can be written as $\eta_i = \gamma(e_{s,t} - \nu_i) + \epsilon$, which is correlated with flu_i because those exposed to the pandemic in utero are also assigned lower SES parental characteristics due to age heaping. This is precisely the bias we aim to overcome.

While age heaping offers an illustrative example, it is worth noting that Brown & Thomas discuss the age heaping phenomenon in their paper and they conclude that age heaping is not an issue. Of course, age heaping is just one potential source of bias. Brown & Thomas use parental characteristics as measured in 1930 or 1920 as controls when estimating outcomes in 1960, 1970, or 1980). An alternative source of bias could be selective mortality (e.g., if low SES children are less likely to survive to be enumerated in the 1960, 1970, or 1980 censuses). If there is any complementarity between this mortality phenomenon and exposure to the pandemic, then Brown & Thomas's proxies for parental characteristics will again be systematically mismeasured. The only way to confidently overcome this problem is to use linked data, which is not possible for Brown & Thomas since they analyze censuses that are not yet publicly available.

7 Conclusion

Almond (2006) provides some of the first evidence in economics in favor of the fetal origins hypothesis by analyzing the 1918 influenza pandemic as a natural experiment. A key assumption of Almond's work is that the 1919 birth cohort would have had similar outcomes to adjacent birth cohorts if the pandemic had never occurred. This assumption is not obvious since other events coincided with the pandemic, most notably the height of WWI. Brown & Thomas (2018) argue that, because servicemen were positively selected from the pool of potential fathers, the 1919 birth cohort had systematically lower SES parents. After accounting for these differences in parental characteristics, they find that the long-run effects of the pandemic disappear.

We estimate the long-run effects of the 1918 influenza pandemic using linked data and city-level variation in influenza exposure. Using linked data allows us to accurately observe parental characteristics with minimal measurement error. Thus, we can directly control for parental differences in the 1919 birth cohort. Second, using city-level data allows us to more accurately measure the local influenza environment and provides us with the necessary variation to include birth cohort fixed effects. Results indicate that the 1918 influenza pandemic reduced educational attainment and the estimates are similar in magnitude to those in Almond (2006).

References

Aizer, Anna, Eli, Shari, Ferrie, Joseph, & Lleras-Muney, Adriana. 2016. The long-run impact of cash transfers to poor families. *American Economic Review*, **106**(4), 935–71.

Almond, Douglas. 2006. Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of Political Economy*, **114**(4), 672–712.

Almond, Douglas, & Mazumder, Bhashkar. 2005. The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data. *American Economic Review*, **95**(2), 258–262.

Bailey, Martha, Cole, Connor, Henderson, Morgan, & Massey, Catherine. 2017. *How Well Do*

Automated Methods Perform in Historical Samples? Evidence from New Ground Truth. Tech. rept. National Bureau of Economic Research.

Barker, David J. 1995. Fetal origins of coronary heart disease. *BMJ: British Medical Journal*, **311**(6998), 171.

Beach, Brian, Ferrie, Joseph, Saavedra, Martin, & Troesken, Werner. 2016. Typhoid fever, water quality, and human capital formation. *Journal of Economic History*, **76**(01), 41–75.

Brown, Ryan, & Thomas, Duncan. 2018. On the long term effects of the 1918 US influenza pandemic. *Unpublished Manuscript*.

Clay, Karen, Lewis, Joshua, & Severnini, Edson. 2015. *Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic*. Tech. rept. National Bureau of Economic Research.

Currie, Janet, & Almond, Douglas. 2011. Human capital development before age five. *Pages 1315–1486 of: Handbook of labor economics*, vol. 4. Elsevier.

Feigenbaum, James J. 2015. Intergenerational mobility during the great depression. *Unpublished working paper*.

Ferrie, Joseph P, & Troesken, Werner. 2008. Water and Chicagos mortality transition, 1850–1925. *Explorations in Economic History*, **45**(1), 1–16.

Frost, Wade Hampton. 1920. Statistics of influenza morbidity: with special reference to certain factors in case incidence and case fatality. *Public Health Reports (1896-1970)*, 584–597.

Goeken, Ronald, Lee, Yu Na, Lynch, Tom, & Magnuson, Diana. 2017. Evaluating the Accuracy of Linked U. S. Census Data: A Household Linking Approach. *Unpublished Manuscript*.

Isen, Adam, Rossin-Slater, Maya, & Walker, W Reed. 2017. Every breath you takeevery dollar youll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy*, **125**(3), 848–902.

Long, Jason, & Ferrie, Joseph. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review*, **103**(4), 1109–37.

Parman, John. 2015a. Childhood health and human capital: New evidence from genetic brothers in arms. *The Journal of Economic History*, **75**(1), 30–64.

Parman, John. 2015b. Childhood health and sibling outcomes: Nurture Reinforcing nature during the 1918 influenza pandemic. *Explorations in Economic History*, **58**, 22–43.

Sanders, Nicholas J. 2012. What doesnt kill you makes you weaker prenatal pollution exposure and educational outcomes. *Journal of Human Resources*, **47**(3), 826–850.

8 Appendix

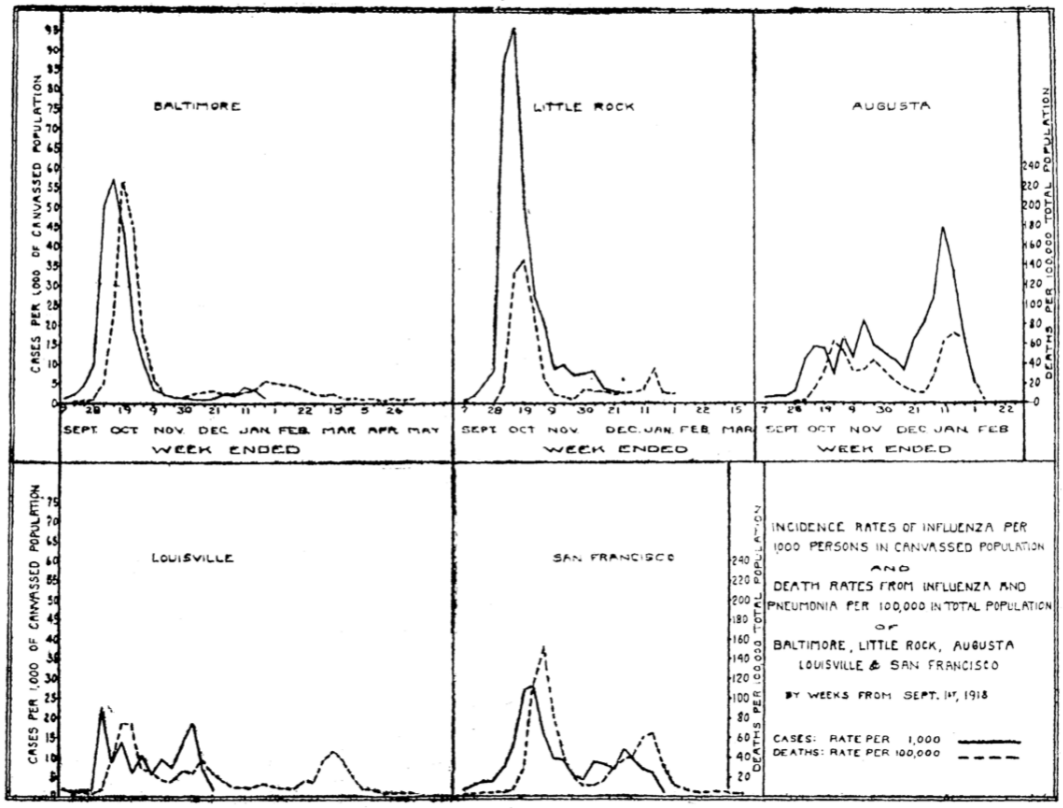
This appendix presents additional figures and tables that were not included in the primary draft.

Figure A.1 reproduces a graph from Frost (1920) in order to illustrate that influenza mortality rates track influenza morbidity rates reasonably well. The slight lag observed in Figure A.1 is not a concern for us because we use annual rather than weekly data to construct our measure of excess influenza.

Figure A.2 plots average excess influenza ratios by year. Note that our sample period is 1900-1917 whereas the 1918-1930 observations are our out of sample predictions. Our model fits the data reasonably very well in the pre pandemic period. The observations in the post pandemic period (1922 and beyond) are noisier, but still close to zero. Further there is no indication of a post pandemic trend in excess influenza.

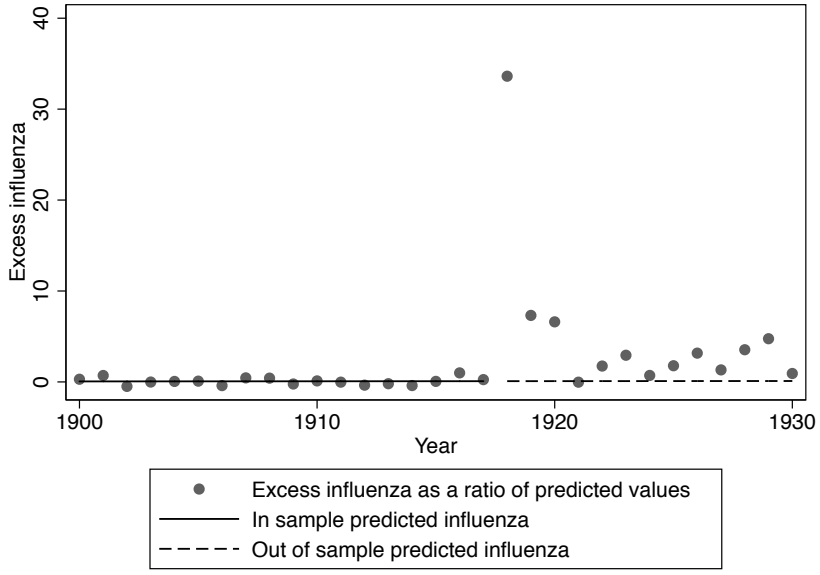
Figure A.3 depicts the spatial variation of excess influenza. The size of each circle relates to the magnitude of exposure. While less clean than the discrete figure in the draft (Figure 1), we continue to see substantial sub-state variation in pandemic exposure.

Figure A.1: Weekly morbidity and mortality for five cities in the fall of 1918



Notes: Figure reprinted from Frost (1920).

Figure A.2: Excess influenza mortality ratio by year



Notes: Excess influenza ratio is calculated by taking actual influenza deaths minus predicted influenza deaths and then dividing by the predicted influenza deaths. City-specific trends in annual ln(influenza mortality) are estimated over the 1900-1917 period.

Figure A.3: Spatial Variation in Flu Intensity

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Excess flu in 1918	36.093	25.025	3.64	179.172	218,662
WWII variables					
Years of schooling	11.478	2.305	8	17	218,662
High school graduate	0.561	0.496	0	1	218,662
Height (inches)	68.284	2.679	50	80	168,317
Weight (pounds)	152.143	22.225	100	300	168,317
Birth year	1916.879	3.535	1909	1922	218,662
Census variables					
Father's age at child's birth	32.066	6.542	18	50	218,662
Father's OccScore	21.712	14.83	0	80	218,662
Father reads and writes	0.949	0.22	0	1	218,662
Father is foreign	0.435	0.496	0	1	218,662
Mother's age at child's birth	28.334	5.899	18	50	218,662
Mother is in labor force	0.044	0.205	0	1	218,662
Mother reads and writes	0.928	0.259	0	1	218,662
Mother is foreign	0.383	0.486	0	1	218,662
Family owns home	0.361	0.48	0	1	218,662
Number of siblings	2.364	1.941	0	9	218,662

Notes: Excess flu in 1918 is calculated by taking actual influenza deaths minus predicted influenza deaths and then dividing by the predicted influenza deaths. See text for more details. WWII variables come from the enlistment records. Sample is restricted to linked males born between 1909 and 1922 in a city for which we have flu data. Census variables are observed in either 1920 for the 1909-1919 birth cohorts or 1930 for the 1920-1922 cohorts. OccScore is the median income for a particular occupation as measured in 1950. This variable is measured in hundreds of 1950 dollars.

Table A.2: In utero exposure and schooling – 1920 links only

	(1)	(2)	(3)	(4)
Panel A: DV is High School Graduate				
Treatment is above median exposure	-0.023***	-0.021**	-0.019***	-0.045
	(0.009)	(0.008)	(0.007)	(0.032)
Continuous treatment (std.)	-0.018***	-0.015***	-0.016***	-0.042**
	(0.004)	(0.005)	(0.004)	(0.020)
Observations	163,263	163,263	163,263	16,219
Panel B: DV is Years of Schooling				
Treatment is above median exposure	-0.091**	-0.077**	-0.031	-0.043
	(0.040)	(0.036)	(0.030)	(0.168)
Continuous treatment (std.)	-0.078***	-0.066***	-0.034	-0.066
	(0.021)	(0.019)	(0.027)	(0.088)
Observations	163,263	163,263	163,263	16,219
Parental controls	N	Y	Y	Y
City-year trends	N	N	Y	Y
Household fixed effects	N	N	N	Y

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (clustered at the city level) in parentheses. Parental controls include: Father's age at time of birth, Mother's age at time of birth, Father's OccScore, indicators for whether the father and mother can read and write, indicators for whether the father and mother are foreign born, an indicator for whether the mother was in the labor force at the time of enumeration, an indicator for whether the family owns their own home, and number of siblings. The continuous treatment specifications interact each city's excess influenza rate with an indicator for being born in 1919. We standardize the excess influenza rate by dividing by the mean excess influenza rate, such that the coefficient can be interpreted as the pandemic's average impact.

Table A.3: In utero exposure and adult health

	(1)	(2)	(3)	(4)
Panel A: DV is Weight				
Treatment is above median exposure	0.157	0.227	0.472	-0.008
	(0.376)	(0.363)	(0.317)	(1.863)
Continuous treatment (std.)	0.159	0.174	0.213	-0.803
	(0.195)	(0.189)	(0.182)	(1.312)
Observations	168,317	168,317	168,317	18,051
Panel B: DV is BMI				
Treatment is above median exposure	0.012	0.021	0.049	0.079
	(0.053)	(0.050)	(0.045)	(0.240)
Continuous treatment (std.)	-0.004	-0.003	0.001	-0.140
	(0.028)	(0.027)	(0.025)	(0.159)
Observations	168,317	168,317	168,317	18,051
Parental controls	N	Y	Y	Y
City-year trends	N	N	Y	Y
Household fixed effects	N	N	N	Y

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (clustered at the city level) in parentheses. Parental controls include: father's age at time of birth, mother's age at time of birth, father's OccScore, indicators for whether the father and mother can read and write, indicators for whether the father and mother are citizens, an indicator for whether the mother was in the labor force at the time of enumeration, an indicator for whether the family owns their home, and number of siblings.

Table A.4: Flu intensity, graduation rates, and parental characteristics in the 1940 census

Panel A: Flu Exposure and High School Graduation

Sample:	All 1918-1920 Cohorts (1)	Residing W/Parents in 1940 (2)	Residing W/Parents in 1940 (3)
Maternal Infection Rate	-0.031***	-0.019**	-0.017**
While In Utero	(0.006)	(0.008)	(0.007)
Parental Controls	N	N	Y
Observations	3,568,851	1,860,609	1,860,609
R-Squared	0.055	0.055	0.160

Panel B: Parental Characteristic Placebo Tests

Dependent Variable:	Pop is HS Grad. (4)	Mom is HS Grad (5)	Mom Age (7)	Pop Age (7)	Mom is Foreign (8)	Mom is Employed (9)	Pop is Employed (10)	Pop's OccScore (11)
Maternal Infection Rate	-0.004	-0.007	-0.065	-0.119	0.020***	0.003	-0.004	0.042
While In Utero	(0.006)	(0.006)	(0.097)	(0.108)	(0.007)	(0.004)	(0.005)	(0.183)
Observations	1,860,609	1,860,609	1,860,609	1,860,609	1,860,609	1,860,609	1,860,609	1,860,609
R-Squared	0.013	0.016	0.007	0.004	0.144	0.009	0.009	0.032

Notes: * p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parentheses. All regressions include year of birth, state of birth, race, and sex fixed effects. The sample is limited by to the 21 states for which we can calculate maternal infection rate (see text).